

0590
1/22



OIPE

#4

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/862,802

DATE: 01/31/2002

TIME: 15:47:34

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\01312002\I862802.raw

ENTERED

3 <110> APPLICANT: Schering-Plough Corporation
5 <120> TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
7 <130> FILE REFERENCE: SF0695B
9 <140> CURRENT APPLICATION NUMBER: US 09/862,802
C--> 10 <141> CURRENT FILING DATE: 2002-01-10
12 <150> PRIOR APPLICATION NUMBER: US 09/111,470
13 <151> PRIOR FILING DATE: 1998-07-08
15 <160> NUMBER OF SEQ ID NOS: 11
17 <170> SOFTWARE: PatentIn version 3.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 1104
21 <212> TYPE: DNA
22 <213> ORGANISM: Unknown
W--> 23 <220> FEATURE:
24 <223> OTHER INFORMATION: mammalian nucleic acid
W--> 26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: 242...952
29 <223> OTHER INFORMATION: protein coding sequence
31 <400> SEQUENCE: 1
32 ttctcactat actggctcctg aggaaagggc ttctgtgaac tgcggttttt agtttttatt 60
34 gtgggttctta gttctcatga gacccctctt gaggatattgt gcctatctgg tgcctctgct 120
36 ctccactagt tgagtgaag gaaggaggta atttaccacc atgtttgggt cctgtttata 180
38 agatgtttta agaaagattt gaaacagatt ttctgaagaa agcagaagct ctcttcccat 240
40 tatgacttcg gaaatcactt atgctgaagt gaggttcaaa aatgaattca agtcctcagg 300
42 catcaacaca gcctcttctg cagcttccaa ggagaggact gcccctctca aaagtaatac 360
44 cggattcccc aagctgcttt gtgcctcact gttgatatatt ttctgtctat tggcaatctc 420
46 attctttatt gcttttgtca ttttctttca aaaatattct cagcttcttg aaaaaaagac 480
48 tacaaaagag ctggttcata caacattgga gtgtgtgaaa aaaaatatgc ccgtggaaga 540
50 gacagcctgg agctgttgcc caaagaattg gaagtcattt agttccaact gctactttat 600
52 ttctactgaa tcagcatctt ggcaagacag tgagaaggac tgtgctagaa tggaggctca 660
54 cctgctggty ataaacactc aagaagagca ggatttcac ttccagaatc tgcaagaaga 720
56 atctgcttat tttgtggggc tctcagatcc agaaggctcag cgacattggc aatgggttga 780
58 tcagacacca tacaatgaaa gttccacatt ctggcatcca cgtgagccca gtgatcccaa 840
60 tgagcgctgc gttgtgctaa attttcgtaa atcacccaaa agatggggct ggaatgatgt 900
62 taattgtctt ggtcctcaaa ggtcagtttg tgagatgatg aagatccact tatgaactga 960
64 acattctcca tgaacaggty gttggattgg tatctgtcat tgtagggata gataataagc 1020
66 tcttcttatt catgtgtaag ggaggccat agaatttagg tgggtctgtca actattctac 1080
68 ttatgagaga attggtctgt acat 1104
71 <210> SEQ ID NO: 2
72 <211> LENGTH: 237
73 <212> TYPE: PRT
74 <213> ORGANISM: Unknown

RAW SEQUENCE LISTING

DATE: 01/31/2002

PATENT APPLICATION: US/09/862,802

TIME: 15:47:34

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\01312002\I862802.raw

76 <220> FEATURE:

77 <223> OTHER INFORMATION: mammalian protein

79 <400> SEQUENCE: 2

81 Met Thr Ser Glu Ile Thr Tyr Ala Glu Val Arg Phe Lys Asn Glu Phe
82 1 5 10 15
85 Lys Ser Ser Gly Ile Asn Thr Ala Ser Ser Ala Ala Ser Lys Glu Arg
86 20 25 30
89 Thr Ala Pro Leu Lys Ser Asn Thr Gly Phe Pro Lys Leu Leu Cys Ala
90 35 40 45
93 Ser Leu Leu Ile Phe Phe Leu Leu Ala Ile Ser Phe Phe Ile Ala
94 50 55 60
97 Phe Val Ile Phe Phe Gln Lys Tyr Ser Gln Leu Leu Glu Lys Lys Thr
98 65 70 75 80
101 Thr Lys Glu Leu Val His Thr Thr Leu Glu Cys Val Lys Lys Asn Met
102 85 90 95
105 Pro Val Glu Glu Thr Ala Trp Ser Cys Cys Pro Lys Asn Trp Lys Ser
106 100 105 110
109 Phe Ser Ser Asn Cys Tyr Phe Ile Ser Thr Glu Ser Ala Ser Trp Gln
110 115 120 125
113 Asp Ser Glu Lys Asp Cys Ala Arg Met Glu Ala His Leu Leu Val Ile
114 130 135 140
117 Asn Thr Gln Glu Glu Gln Asp Phe Ile Phe Gln Asn Leu Gln Glu Glu
118 145 150 155 160
121 Ser Ala Tyr Phe Val Gly Leu Ser Asp Pro Glu Gly Gln Arg His Trp
122 165 170 175
125 Gln Trp Val Asp Gln Thr Pro Tyr Asn Glu Ser Ser Thr Phe Trp His
126 180 185 190
129 Pro Arg Glu Pro Ser Asp Pro Asn Glu Arg Cys Val Val Leu Asn Phe
130 195 200 205
133 Arg Lys Ser Pro Lys Arg Trp Gly Trp Asn Asp Val Asn Cys Leu Gly
134 210 215 220
137 Pro Gln Arg Ser Val Cys Glu Met Met Lys Ile His Leu
138 225 230 235

141 <210> SEQ ID NO: 3

142 <211> LENGTH: 1458

143 <212> TYPE: DNA

144 <213> ORGANISM: Unknown

146 <220> FEATURE:

147 <223> OTHER INFORMATION: mammalian nucleic acid

149 <220> FEATURE:

150 <221> NAME/KEY: CDS

151 <222> LOCATION: 257...1204

152 <223> OTHER INFORMATION: protein coding sequence

154 <220> FEATURE:

155 <221> NAME/KEY: variation

156 <222> LOCATION: 608...673

157 <223> OTHER INFORMATION: short form lacks these nucleotides

159 <400> SEQUENCE: 3

160 gttgaggaga tgggatgtcc cagatgatag ggctcctggg atttcagacc caagaccagc 60

RAW SEQUENCE LISTING

DATE: 01/31/2002

PATENT APPLICATION: US/09/862,802

TIME: 15:47:34

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\01312002\I862802.raw

```
162 aggactccag tcacctctac cccagctctc caggacacag cgctcccaac tctgagtgc 120
164 gtcccacctc tggtccttgc agcacaacca acgtgggaat cacaccctcc agacctccca 180
166 cagctccacc ccagactggg cgccggccct gcctccattt cagctgtgac aacctcagag 240
168 ccgtgttggc ccaagcatga caaggacgta tgaaaacttc cagtacttgg agaataaggt 300
170 gaaagtccag gggtttaaaa atggggccact tcctctccag tccctcctgc agcgtctccg 360
172 ctctgggccc tgccatctcc tgctgtccct gggcctcggc ctgctgctgc tggatcatcat 420
174 ctgtgtggtt ggattccaaa attccaaatt tcagagggac ctggtgaccc tgagaacaga 480
176 ttttagcaac ttcacctcaa acactgtggc ggagatccag gcaactgactt cccagggcag 540
178 cagcttggaa gaaacgatag catctctgaa agctgagggtg gaggggtttca agcaggaacg 600
180 gcaggcaggg gtatctgagc tccaggaaca cactacgcag aaggcacacc taggccactg 660
182 tcccactgc ccatctgtgt gtgtcccagt tcattctgaa atgctcctgc gagtccagca 720
184 gctgggtgcaa gacctgaaga aactgacctg ccagggtggct actctcaaca acaatgcctc 780
186 cactgaaggg acctgctgcc ccgtcaactg ggtggagcac caagacagct gctactgggtt 840
188 ctctcactct gggatgtcct gggccgaggg tgagaagtac tgccagctga agaacgcccc 900
190 cctgggtggc atcaactcca gggaggagca gaattttgtc cagaaatatc taggctccgc 960
192 atacacctgg atgggcctca gtgacctga aggagcctgg aagtgggtgg atggaacaga 1020
194 ctatgcgacc ggcttccaga actggaagcc aggccagcca gacgactggc aggggacagg 1080
196 gctgggtgga ggcgaggact gtgctcactt ccatccagac ggcaggtgga atgacgacgt 1140
198 ctgccagagg ccctaccact gggctctgca ggcctggcctg ggtcagacca gccaggagag 1200
200 tcaactgagc gcctttgggt ggaccaccgg gccacagaaa tggcgggtggg aggaggactc 1260
202 ttctcacgac ctctctgcaa gaccgctctg ggagagaaat aagcactggg agattggaag 1320
204 cactgctaac attttgaatt tttttctctt taattttaaa aagatgggtat agtgttctta 1380
206 agcttttatt ttttttccaa cttttgaaag tcaacttcat gaagggtataa tttttacata 1440
208 ataaaaatgc actcattt 1458
211 <210> SEQ ID NO: 4
212 <211> LENGTH: 316
213 <212> TYPE: PRT
214 <213> ORGANISM: Unknown
216 <220> FEATURE:
217 <223> OTHER INFORMATION: mammalian protein
219 <400> SEQUENCE: 4
221 Met Thr Arg Thr Tyr Glu Asn Phe Gln Tyr Leu Glu Asn Lys Val Lys
222 1 5 10 15
225 Val Gln Gly Phe Lys Asn Gly Pro Leu Pro Leu Gln Ser Leu Leu Gln
226 20 25 30
229 Arg Leu Arg Ser Gly Pro Cys His Leu Leu Leu Ser Leu Gly Leu Gly
230 35 40 45
233 Leu Leu Leu Leu Val Ile Ile Cys Val Val Gly Phe Gln Asn Ser Lys
234 50 55 60
237 Phe Gln Arg Asp Leu Val Thr Leu Arg Thr Asp Phe Ser Asn Phe Thr
238 65 70 75 80
241 Ser Asn Thr Val Ala Glu Ile Gln Ala Leu Thr Ser Gln Gly Ser Ser
242 85 90 95
245 Leu Glu Glu Thr Ile Ala Ser Leu Lys Ala Glu Val Glu Gly Phe Lys
246 100 105 110
249 Gln Glu Arg Gln Ala Gly Val Ser Glu Leu Gln Glu His Thr Thr Gln
250 115 120 125
253 Lys Ala His Leu Gly His Cys Pro His Cys Pro Ser Val Cys Val Pro
254 130 135 140
```

RAW SEQUENCE LISTING

DATE: 01/31/2002

PATENT APPLICATION: US/09/862,802

TIME: 15:47:34

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\01312002\I862802.raw

```

257 Val His Ser Glu Met Leu Leu Arg Val Gln Gln Leu Val Gln Asp Leu
258 145 150 155 160
261 Lys Lys Leu Thr Cys Gln Val Ala Thr Leu Asn Asn Asn Ala Ser Thr
262 165 170 175
265 Glu Gly Thr Cys Cys Pro Val Asn Trp Val Glu His Gln Asp Ser Cys
266 180 185 190
269 Tyr Trp Phe Ser His Ser Gly Met Ser Trp Ala Glu Ala Glu Lys Tyr
270 195 200 205
273 Cys Gln Leu Lys Asn Ala His Leu Val Val Ile Asn Ser Arg Glu Glu
274 210 215 220
277 Gln Asn Phe Val Gln Lys Tyr Leu Gly Ser Ala Tyr Thr Trp Met Gly
278 225 230 235 240
281 Leu Ser Asp Pro Glu Gly Ala Trp Lys Trp Val Asp Gly Thr Asp Tyr
282 245 250 255
285 Ala Thr Gly Phe Gln Asn Trp Lys Pro Gly Gln Pro Asp Asp Trp Gln
286 260 265 270
289 Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala His Phe His Pro Asp
290 275 280 285
293 Gly Arg Trp Asn Asp Asp Val Cys Gln Arg Pro Tyr His Trp Val Cys
294 290 295 300
297 Glu Ala Gly Leu Gly Gln Thr Ser Gln Glu Ser His
298 305 310 315
301 <210> SEQ ID NO: 5
302 <211> LENGTH: 291
303 <212> TYPE: PRT
304 <213> ORGANISM: Unknown
306 <220> FEATURE:
307 <223> OTHER INFORMATION: mammalian protein
309 <400> SEQUENCE: 5
311 Met Thr Lys Glu Tyr Gln Asp Leu Gln His Leu Asp Asn Glu Glu Ser
312 1 5 10 15
315 Asp His His Gln Leu Arg Lys Gly Pro Pro Pro Pro Gln Pro Leu Leu
316 20 25 30
319 Gln Arg Leu Cys Ser Gly Pro Arg Leu Leu Leu Leu Ser Leu Gly Leu
320 35 40 45
323 Ser Leu Leu Leu Leu Val Val Val Cys Val Ile Gly Ser Gln Asn Ser
324 50 55 60
327 Gln Leu Gln Glu Glu Leu Arg Gly Leu Arg Glu Thr Phe Ser Asn Phe
328 65 70 75 80
331 Thr Ala Ser Thr Glu Ala Gln Val Lys Gly Leu Ser Thr Gln Gly Gly
332 85 90 95
335 Asn Val Gly Arg Lys Met Lys Ser Leu Glu Ser Gln Leu Glu Lys Gln
336 100 105 110
339 Gln Lys Asp Leu Ser Glu Asp His Ser Ser Leu Leu Leu His Val Lys
340 115 120 125
343 Gln Phe Val Ser Asp Leu Arg Ser Leu Ser Cys Gln Met Ala Ala Leu
344 130 135 140
347 Gln Gly Asn Gly Ser Glu Arg Thr Cys Cys Pro Val Asn Trp Val Glu
348 145 150 155 160

```

RAW SEQUENCE LISTING

DATE: 01/31/2002

PATENT APPLICATION: US/09/862,802

TIME: 15:47:34

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\01312002\I862802.raw

```

351 His Glu Arg Ser Cys Tyr Trp Phe Ser Arg Ser Gly Lys Ala Trp Ala
352                               165                               170                               175
355 Asp Ala Asp Asn Tyr Cys Arg Leu Glu Asp Ala His Leu Val Val Val
356                               180                               185                               190
359 Thr Ser Trp Glu Glu Gln Lys Phe Val Gln His His Ile Gly Pro Val
360                               195                               200                               205
363 Asn Thr Trp Met Gly Leu His Asp Gln Asn Gly Pro Trp Lys Trp Val
364                               210                               215                               220
367 Asp Gly Thr Asp Tyr Glu Thr Gly Phe Lys Asn Trp Arg Pro Glu Gln
368 225                               230                               235                               240
371 Pro Asp Asp Trp Tyr Gly His Gly Leu Gly Gly Gly Glu Asp Cys Ala
372                               245                               250                               255
375 His Phe Thr Asp Asp Gly Arg Trp Asn Asp Asp Val Cys Gln Arg Pro
376                               260                               265                               270
379 Tyr Arg Trp Val Cys Glu Thr Glu Leu Asp Lys Ala Ser Gln Glu Pro
380                               275                               280                               285
383 Pro Leu Leu
384       290
387 <210> SEQ ID NO: 6
388 <211> LENGTH: 287
389 <212> TYPE: PRT
390 <213> ORGANISM: Unknown
392 <220> FEATURE:
393 <223> OTHER INFORMATION: mammalian protein
395 <400> SEQUENCE: 6
397 Met Ala Lys Asp Phe Gln Asp Ile Gln Gln Leu Ser Ser Glu Glu Asn
398 1                               5                               10                               15
401 Asp His Pro Phe His Gln Gly Pro Pro Pro Ala Gln Pro Leu Ala Gln
402                               20                               25                               30
405 Arg Leu Cys Ser Met Val Cys Phe Ser Leu Leu Ala Leu Ser Phe Asn
406                               35                               40                               45
409 Ile Leu Leu Leu Val Val Ile Cys Val Thr Gly Ser Gln Ser Ala Gln
410                               50                               55                               60
413 Leu Gln Ala Glu Leu Arg Ser Leu Lys Glu Ala Phe Ser Asn Phe Ser
414 65                               70                               75                               80
417 Ser Ser Thr Leu Thr Glu Val Gln Ala Ile Ser Thr His Gly Gly Ser
418                               85                               90                               95
421 Val Gly Asp Lys Ile Thr Ser Leu Gly Ala Lys Leu Glu Lys Gln Gln
422                               100                              105                              110
425 Gln Asp Leu Lys Ala Asp His Asp Ala Leu Leu Phe His Leu Lys His
426                               115                              120                              125
429 Phe Pro Val Asp Leu Arg Phe Val Ala Cys Gln Met Glu Leu Leu His
430                               130                              135                              140
433 Ser Asn Gly Ser Gln Arg Thr Cys Cys Pro Val Asn Trp Val Glu His
434 145                              150                              155                              160
437 Gln Gly Ser Cys Tyr Trp Phe Ser His Ser Gly Lys Ala Trp Ala Glu
438                               165                              170                              175
441 Ala Glu Lys Tyr Cys Gln Leu Glu Asn Ala His Leu Val Val Ile Asn
442                               180                              185                              190

```

VERIFICATION SUMMARY

DATE: 01/31/2002

PATENT APPLICATION: US/09/862,802

TIME: 15:47:35

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\01312002\I862802.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:23 M:283 W: Missing Blank Line separator, <220> field identifier

L:611 M:283 W: Missing Blank Line separator, <220> field identifier